

# EUROPEAN PATENT OFFICE

## Patent Abstracts of Japan

PUBLICATION NUMBER : 02131300  
PUBLICATION DATE : 21-05-90

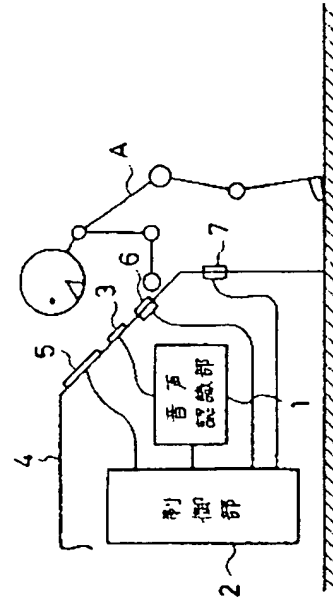
APPLICATION DATE : 11-11-88  
APPLICATION NUMBER : 63285504

APPLICANT : TOSHIBA CORP;

INVENTOR : SUGI NOBUO;

INT.CL. : G10L 3/00 G06F 3/16

TITLE : VOICE RECOGNIZING DEVICE



ABSTRACT : PURPOSE: To securely input a voice that a speaker voices without omitting its head part and to enable voice recognizing by starting voice recognizing processing for a voice which is inputted from a microphone when the approach state of the speaker is detected by an auxiliary proximity sensor.

CONSTITUTION: The auxiliary proximity sensor 7 is provided to a housing 4 and the proximity sensor 6 detects the speaker A being close to the housing 4 before the proximity sensor 6. According to the detection result of this auxiliary proximity sensor 7, a proximity sensor 6 before detecting the face of the speaker A approaching a microphone 3 supplies a recognition start command from a control circuit 2 to a voice recognizing part 1, thereby starting recognition processing. Consequently, even when the voicing start timing of the speaker A is too early, the speaker's voice is securely inputted and voice recognizing is performed.

COPYRIGHT: (C)1990,JPO&Japio

⑩ 日本国特許庁(JP)

⑪ 特許出願公開

## ⑫ 公開特許公報(A) 平2-131300

⑬ Int. Cl.<sup>5</sup>

識別記号

庁内整理番号

⑭ 公開 平成2年(1990)5月21日

G 10 L 3/00

3 0 1 Z

8842-5D

G 06 F 3/16

3 2 0 F

8323-5B

G 10 L 3/00

3 0 1 A

8842-5D

審査請求 未請求 請求項の数 1 (全5頁)

⑮ 発明の名称 音声認識装置

⑯ 特 願 昭63-285504

⑰ 出 願 昭63(1988)11月11日

⑱ 発 明 者 杉 伸 夫 神奈川県川崎市幸区柳町70番地 株式会社東芝柳町工場内

⑲ 出 願 人 株式会社東芝 神奈川県川崎市幸区堀川町72番地

⑳ 代 理 人 弁理士 鈴江 武彦 外2名

## 明 題 書

## 1. 発明の名称

音声認識装置

## 2. 特許請求の範囲

筐体に設けられたマイクロフォンにより発話者が発声した音声を入力して音声認識するに際し、上記マイクロフォンの近傍位置に設けられた近接センサにより前記発話者の顔と上記マイクロフォンとの近接状態を検出し、この近接状態に応じて発話者に音声入力に関するメッセージを出力するようにした音声認識装置において、

前記筐体に発話者が近付いたことを検出する為の補助近接センサを設け、この補助近接センサにより発話者の近接状態が検出されたとき、前記マイクロフォンから入力される音声に対する音声認識処理を開始させることを特徴とする音声認識装置。

## 3. 発明の詳細な説明

## 〔発明の目的〕

(産業上の利用分野)

本発明は発話者の音声を確実に収集して音声認識することのできる音声認識装置に関する。

(従来の技術)

近時、種々の機器に音声認識装置を組込み、発話者が発声した命令等の音声をマイクロホンを通じて取込んでその入力音声を認識処理し、その認識結果に従って種々の情報処理を実行することが種々試みられている。例えばワードプロセッサにおける文書情報入力手段や、プラント等における制御命令入力手段、更には種々の販売機における情報入力手段等として音声認識装置の利用が注目されている。

ところでこの種の音声入力装置にあっては、周囲雑音等の悪影響を受けることなしに情報入力目的とする音声だけを確実に入力し、これを音声認識することがその認識性能を高める上で重要な課題となる。しかして音声認識処理は、一般にその

## 特開平2-131300 (2)

制御部（ホスト計算機）からの認識開始コマンドを受けて開始されるが、実際のシステムでは音声認識処理の開始のタイミングを適切に設定することが非常に難しい。そこで従来では、一般的に発話者が発声した音声を入力する為のマイクロフォンの近傍位置に超音波センサ等の近接センサを設け、この近接センサにより発話者の顔（口元）とマイクロフォンとの近接状態を検出し、例えば発話者の顔がマイクロフォンに十分に近付いたことが検出されたときに認識開始コマンドを発し、その後、音声の入力開始を指示する旨のメッセージを出力するようにしている。またこの近接センサにて発話者の顔がマイクロフォンから離れていることが検出されたときには、顔（口）をマイクロフォンに近付けて発声する旨をメッセージ出力し、良好な音声入力状態を確保するものとなっている。

ところがこのようにして近接センサを用いてマイクロフォンと発話者の顔との近接状態を検出して音声認識処理の動作を制御し、発話者に対して音声入力に関するメッセージを出力して音声入力

のタイミングやその入力姿勢を指示するにも拘らず、発話者によってはマイクロフォンに十分に近付かない状態で、或いはマイクロフォンに十分に近付いたとしても音声入力開始のメッセージが出力される前に発話を開始してしまうことが多々ある。このような場合には、音声認識の開始コマンドが与えられて音声認識処理が開始される以前に音声入力がなされることになるので、音声認識装置にはその先頭部分が欠落した状態で音声が入力されることになる。この結果、誤認識や認識リジェクトが発生し、発話者に音声の再入力を促すことが必要となり、発話者のみならず音声認識装置にとっても従に負担が増える。

特に発話の開始が発生要求のタイミングより僅かに早いだけの場合には、発話者自身に発声タイミングが早いと云う意識がない為、音声の再入力時にも同様にしてその先頭部分が欠落してしまうことが多いと云う問題がある。

このような不具合を解消する為、近接センサの感度を高めておき、発話者がマイクロフォンに

近付いてくる途中のタイミングで認識開始コマンドを与えることが考えられている。しかしこのような対策を講じた場合、発話者がマイクロフォンに十分に近付かない状態で発話開始することについての対策を講じることができなくなり、誤認識や認識リジェクトが増える原因となる。また上述した感度の高い近接センサを用い、発話者が近づくことを検出した後、一定時間を経て発話要求を発することも考えられているが、そのタイミング設定が非常に難しく、また発話者がマイクロフォンに十分近付いたか否かを検出することができないと云う問題がある。

## （発明が解決しようとする課題）

このように従来にあつては、発話者がマイクロフォンに十分に近付いたことを検出し、先頭部分の欠落を生じることなく上記発話者からの音声を取込んで認識する等で、認識処理開始のタイミング等を適切に設定することが非常に困難であつた。

本発明はこのような事情を考慮してなされたも

ので、その目的とするところは、発話者が発声する音声とその先頭部分の欠落を生じることなく確實に入力して音声認識することのできる音声認識装置を提供することにある。

## （課題を解決するための手段）

本発明は、筐体に設けられたマイクロフォンにより発話者が発声した音声を入力して音声認識する際、上記マイクロフォンの近傍位置に設けられた近接センサにより前記発話者の顔と上記マイクロフォンとの近接状態を検出し、この近接状態に応じて発話者に音声入力に関するメッセージを出力するようにした音声認識装置に係り、

前記近接センサとは別に前記筐体に発話者が近付いたことを検出する為の補助近接センサを設け、この補助近接センサにより発話者の近接状態が検出されたとき、前記マイクロフォンから入力される音声に対する音声認識処理を開始させるようにしたことを特徴とするものである。

つまり近接センサによる検出結果に従って音声入力のタイミング等を促すに先立って、補助近接

## 特開平2-131300 (3)

センサによる検出結果に従って音声認識処理を開始させておくようにしたことを特徴とするものである。

## (作 用)

本発明によれば、近接センサによる検出結果に従って音声入力タイミング等を促すに先立って補助近接センサにより発話者の近接を検出し、この検出結果に従って音声認識処理を開始させておくので、発話者の発声開始タイミングが早い場合であっても、その先頭部分の欠落を生じることなく発話音声を確実に取込んで音声認識することが可能となる。

この結果、非常に簡易にして効果的に誤認識や認識リジェクトの発生を抑えてその認識効率を高めることが可能となる。

## (実施例)

以下、図面を参照して本発明の一実施例に係る音声認識装置につき説明する。

第1図は実施例装置の要部概略構成図であり、第2図はその制御処理シーケンスを示す図である。

示出力する。この表示器5に表示されるメッセージは、例えば

「音声入力を開始して下さい。」

「しっとマイクロフォンに近付いて

発声して下さい。」

等のメッセージからなる。尚、音声合成装置等を用いてこれらのメッセージを音声出力することも勿論可能である。

しかし本装置が特徴とするところは、上述した構成に加えて前記筐体4に補助近接センサ(第2の近接センサ)7を設け、前記近接センサ(第1の近接センサ)6により発話者Aの顔(口元)がマイクロフォン3に近付いたことが検出される以前に、その発話者Aが本装置(筐体4)に近接したことを検出するようにした点にある。そしてこの補助近接センサ(第2の近接センサ)7による検出結果に従い、近接センサ(第1の近接センサ)8により発話者Aの顔(口元)がマイクロフォン3に近付いたことが検出されるに先立って、前記制御部2から音声認識部1に対して認識開始

第1図において1は各種認識方式を採用して構成される音声認識部である。この音声認識部1は制御部2からの認識開始コマンドを受けて音声認識処理を開始し、マイクロフォン3から入力される音声を認識処理し、その認識結果を出力する。

しかしこれらの音声認識部1および制御部2は、一般的には各種情報処理装置本体の筐体4内に収められて設けられ、前記マイクロフォン3は上記筐体4の所定操作パネル面に設けられる。

このマイクロフォン3を設けた操作パネル面には、発話者Aに対して音声入力に関する種々のメッセージを表示する表示器5や、上記マイクロフォン3への発話者Aの顔(口元)の近接を検出する近接センサ(第1の近接センサ)6が上記マイクロフォン3の近傍位置に設けられている。前記制御部2は、基本的にはこの近接センサ(第1の近接センサ)6により発話者Aの顔(口元)のマイクロフォン3に対する近接状態を検出し、この検出結果(近接状態)に応じて前記表示器5を用いて音声入力に関する種々のメッセージを認

知コマンドを与え、認識処理を開始させるようにした点にある。

このような本装置の特徴ある動作を図2図を参照して説明すると、先ず補助近接センサ(第2の近接センサ)7にて発話者Aが本装置(筐体4)に所定の距離範囲内に近接したか否かを検出し、その近接状態を監視する(ステップa)。そして発話者Aが本装置(筐体4)に所定の距離範囲内(例えば30cm程度)に近接したことが上記補助近接センサ(第2の近接センサ)7にて検出されたとき、制御部2から音声認識部1に対して認識開始コマンドを出力し(ステップb)、音声認識部1にその音声認識処理動作を開始させる(ステップc)。

この状態で次に前記近接センサ(第1の近接センサ)8にて発話者Aの顔(口元)がマイクロフォン3に対して所定の距離範囲内(例えば10cm程度)に近接したか否かを調べる(ステップd)。尚、この近接センサ(第1の近接センサ)6による発話者Aの顔(口元)のマイクロフォン3に対

## 特開平2-131300 (4)

する近接検出は、前記補助近接センサ（第2の近接センサ）7 による発話者Aの近接検出タイミングから所定時間内に発話者Aの顔（口元）がマイクロフォン3 に近接したか否かを監視しながら行なわれ（ステップe）、所定時間内にその近接が検出されない場合には（タイムアウト時）、前記表示器5 を用いてマイクロフォンに対して顔を近付ける旨のメッセージを出力する（ステップf）。この所定時間は、一般的には2〜3秒程度に設定すれば十分である。

このような制御手続を経て発話者Aの顔がマイクロフォン3 に十分に近付いたことが検出されたとき（ステップd）、前記表示器5 を用いて音声入力を開始を促す旨のメッセージを出力する（ステップg）。そしてこのメッセージに従って発声される音声を入力処理し（ステップh）、その入力音声について前記音声認識部1 にて認識処理を実行させる（ステップi）。そしてその認識結果を前記音声認識部1 から制御部2 に対して転送し（ステップj）、適宜発話者に対してその認

識結果の確認表示を行なう等して、一連の音声認識処理制御を終了する。

かくしてこのように構成され、且つ上述したように制御処理動作する本装置によれば、発話者Aが近付いた時点で音声認識部1 における認識処理を開始させ、その後、発話者Aがマイクロフォン3 に近付いた時点で音声入力の開始を促す旨のメッセージを出力するので、既に発話者Aの発声開始タイミングがそのメッセージ表示のタイミングよりも段々に早くても、その先頭部分の欠落を招くことなく入力音声を音声認識部1 に取込んで認識処理を行なわせることが可能となり、発話者Aが発声した音声を確実に認識することが可能となる。従って先頭部分の欠落に起因する誤認識や認識リジェクトの発生をなくし、その認識効率を高めることが可能となる。また先頭部分の欠落に起因する誤認識や認識リジェクトがない分だけ発話者Aに対して音声の入力力を促す頻度が減るので、発話者に対する負担の軽減を図り得ると共に、装置自体にとってもその処理負担の軽減を図ることが

可能となる。

また実施例装置によれば、発話者Aがマイクロフォン3 に近接していない状態が検出されたとき、適宜注意を促すメッセージを出力することができるので、これによって安定な音声入力状況を効果的に確保することが可能となる。

つまり2つの近接センサ8,7 によりその機能を分担して担わせ、これらの近接センサ8,7 による状態検出結果に従って音声認識処理を制御するので、発話者Aが発声した音声を確実に認識することが可能となる。

尚、本発明は上述した実施例に限定されるものではない。例えば発話者の近接距離をアナログ的に検出を以て検出し得る近接センサ（距離測定器）を用いることにより、上述した2つの機能を実質的に1つのセンサに担わせることも可能である。またここでは音声認識部1 とその制御部2 とを筐体4 内に収容した構造のものについて説明したが、マイクロフォン3 や近接センサ8,7 とその筐体を分離して設けることも勿論可能である。またここ

で採用される音声認識方式自体については特に限定されないことも言うまでもなく、要はその要旨を逸脱しない範囲で種々変形して実施することができる。

## 【発明の効果】

以上説明したように本発明によれば、発話者がマイクロフォンに近付くことが検出される先立って音声認識処理を開始させるようにその制御が行なわれるので、発話者の発声タイミングが早い場合であってもその先頭部分の欠落を生じることなく、その発話音声を確実に認識することができる等の実用上多大なる効果が奏せられる。

## 4. 図面の簡単な説明

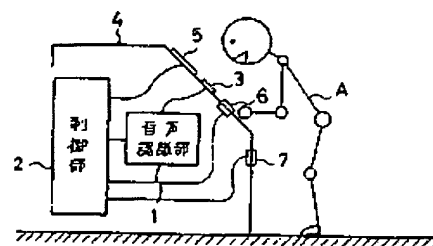
図は本発明の一実施例に係る音声認識装置につき示すもので、第1図は実施例装置の要部概略構成図、第2図はその制御処理シーケンスを示す図である。

1…音声認識部、2…制御部、3…マイクロフォン、4…筐体、5…表示器、6…近接センサ（第1の近接センサ）、7…補助近接センサ（第

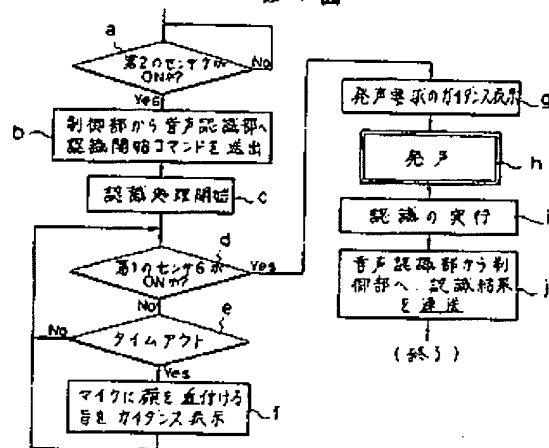
特開平2-131300 (5)

2の近接センサ)、A…電話者。

出版人代理人 井理士 崎江武彦



第1図



第2図

特開平2-131300

【公報種別】特許法第17条の2の規定による補正の掲載  
 【部門区分】第6部門第2区分  
 【発行日】平成9年(1997)1月10日

【公開番号】特開平2-131300  
 【公開日】平成2年(1990)5月21日  
 【年通号数】公開特許公報2-1313  
 【出願番号】特願昭63-285504  
 【国際特許分類第6版】

G10L 3/00 511  
 G06F 3/16 320  
 G10L 3/00 571

【F I】

G10L 3/00 511 9379-SH  
 G06F 3/16 320 F 9172-5E  
 G10L 3/00 571 K 9379-SH

特 許 公 報 特 許 庁

平成 9 年 1 月 10 日

特許庁長官 市川 信二 殿

## 1. 事件の表示

特願昭63-285504号

## 2. 発明の名称

音声認識装置

## 3. 発明をする者

発明者の氏名 特許出願人

(307) 株式会社 東芝

## 4. 代理人

東京都千代田区森が国3丁目7番2号  
 丸 型 内 外 開 特 許 事 務 所 内  
 〒100 電話03(3502)3111(大代表)  
 (5847) 伊藤士 裕 正 氏 啓

## 5. 自 発 補 正

## 6. 補正の手段

説明書

## 7. 補正の内容

特許請求の範囲を別紙の通り訂正する。

## 2. 特許請求の範囲

図体に設けられたマイクローホンにより発話者が発した音声を取り込んで、  
 その入力音声を認識する音声認識装置において、

前記発話者が第1の距離に近接したことを検出する第1の近接センサと、

前記発話者が第2の距離に近接したことを検出する第2の近接センサと、

前記第1の近接センサにて、前記発話者の顔が、第1の距離に近接したことが  
 検出された時、前記発話者に対して音声入力に関するメッセージを出力する手段  
 と、

前記第2の近接センサにて、前記発話者が、第2の距離に近接したことが検出  
 された時、前記マイクローホンから入力される音口に対する音声認識を開始する  
 手段とを具備したことを特徴とする音声認識装置。

出願人代理人 弁護士 松 井 純 氏

